British Journal of Educational Psychology (2015), 85, 33–46

The British Psychological Society

www.wileyonlinelibrary.com

The relation between Assessment for Learning and elementary students' cognitive and metacognitive strategy use

Diana Baas[1]*, Jos Castelijns[2], Marjan Vermeulen[3], Rob Martens[3] and Mien Segers[1]

[1]Educational Research and Development, Maastricht University, The Netherlands,
[2]De Kempel Institute for Teacher Education, Helmond, The Netherlands
[3]Welten Institute, Open University, Heerlen, The Netherlands

Background. Assessment for Learning (AfL) is believed to create a rich learning environment in which students develop their cognitive and metacognitive strategies. Monitoring student growth and providing scaffolds that shed light on the next step in the learning process are hypothesized to be essential elements of AfL that enhance cognitive and metacognitive strategies. However, empirical evidence for the relation between AfL and students' strategy use is scarce.

Aim. This study investigates the relation between AfL and elementary school students' use of cognitive and metacognitive strategies.

Sample. The sample comprised 528 grade four to six students (9- to 12-year-olds) from seven Dutch elementary schools.

Methods. Students' perceptions of AfL and their cognitive and metacognitive strategy use were measured by means of questionnaires. Structural equation modelling was used to investigate the relations among the variables.

Results. The results reveal that monitoring activities that provide students an understanding of where they are in their learning process predict students' task orientation and planning. Scaffolding activities that support students in taking the next step in their learning are positively related to the use of both surface and deep-level learning strategies and the extent to which they evaluate their learning process after performing tasks.

Conclusions. The results underline the importance of assessment practices in ceding responsibility to students in taking control of their own learning.

An important challenge for education lies in supporting students in learning how to learn. Schools are stimulating students in developing habits and skills that enable them to learn independently throughout their academic career and beyond (Boekaerts, 1999). In this respect, particular relevance is given to the concept of self-regulated learning (Boekaerts, 1997; Paris & Paris, 2001; Pintrich, 2004; Schunk, 1990; Winne, 1996; Zimmerman, 1990). Assessment is increasingly considered a tool to equip students with a repertoire of cognitive and metacognitive strategies to become more effective learners who are able to

*Correspondence should be addressed to Diana Baas, Department of Educational Research and Development, Faculty of Business and Economics, Maastricht University, PO Box 616, Maastricht, 6200 MD, The Netherlands (email: d.baas@maastrichtuniversity.nl).

DOI:10.1111/bjep.12058

self-regulate their learning (Black & Wiliam, 2009; Clark, 2012). In particular, the beneficial impact of Assessment for Learning (AfL) on a range of aspects of student learning has been an inspiration for many researchers (e.g., ARG, 2002; Black & Wiliam, 1998a; Wiliam, Lee, Harrison, & Black, 2004). AfL is the process of collecting information about student learning from a wide variety of assessment practices and using this information to modify teaching and learning in order to better meet students' needs (ARG, 2002; Wiliam, 2011). While teachers play a role in the assessment process, the ultimate goal is to stimulate students in becoming increasingly autonomous in their learning process (Klenowski, 2009). Therefore, AfL provides students a rich learning environment in which they can take responsibility for their own learning and develop a range of cognitive and metacognitive strategies to achieve this.

Although numerous scholars (e.g., Birenbaum, Kimron, Shilton, & Shahaf-Barzilay, 2009; Black & Wiliam, 2009; Clark, 2012; Nicol & Macfarlane-Dick, 2006) argue that AfL promotes student learning, empirical evidence is not undisputed (Bennett, 2011; Kingston & Nash, 2011). While Black and Wiliam (1998a) put forward evidence for the effect of AfL practices that seems convincing; several scholars (e.g., Bennett, 2011; Dunn & Mulvenon, 2009; Kingston & Nash, 2011) have highlighted that many studies on which these efficacy claims are based are flawed and therefore difficult to interpret. In a meta-analysis on the impact of AfL, Kingston and Nash (2011) conclude that there is a wide variation in the impact, ranging from small to moderate effects. Also, studies have been reported that failed to find a significant effect (e.g., Wiliam *et al.*, 2004; Yin *et al.*, 2008). Interpreting the findings of studies on the effectiveness of AfL is difficult as many of these studies focus on enhanced student achievement as an outcome measure. This entails the risk that a beneficial effect of AfL merely reflects effective teaching to the test. Consequently, empirical support for the beneficial impact of AfL on student learning is needed. In particular, the relation between AfL and cognitive and metacognitive strategy use has received scarce empirical attention. This is surprising while numerous claims have been made about this relation. Therefore, this study aims at narrowing this gap and investigates the relation between AfL and elementary school students' cognitive and metacognitive strategies.

Cognitive and metacognitive strategies as part of self-regulated learning

Students' ability to take responsibility and control for their learning is described as self-regulation. The concept of self-regulated learning refers to 'self-generated thoughts, feelings and actions that are planned and cyclically adapted to the attainment of personal goals' (Zimmerman, 2000, p. 14). It involves the use of motivational strategies, cognitive strategies, and metacognitive strategies. In terms of motivation, self-regulated learners have an intrinsic interest in learning and report high levels of self-efficacy. Cognitively, self-regulated learners optimize their learning environment by selecting learning strategies and structuring their environment. The metacognitive aspect of self-regulated learning refers to students' ability to plan and organize learning activities, set goals, and evaluate their learning at various points during the process (Zimmerman, 1990). In this study, we only focus on the cognitive and metacognitive aspects of self-regulated learning.

The ability to self-regulate learning is generally conceptualized in process terms instead of as a personal trait (Boekaerts & Corno, 2005; Zimmerman, 2000). Zimmerman (2000, 2008) separates the complex process of self-regulation into three sequential phases that students go through when they perform a task: Forethought, performance, and self-reflection. In the forethought phase, students analyse tasks, set goals for themselves, and plan their learning activities. While performing the task, learners need to focus their attention on the task and control the execution of learning strategies. A distinction is made between surface learning strategies and deep-level learning strategies (Marton & Säljö, 1997). The use of surface learning strategies is associated with memorizing learning material and getting a basic understanding of the learning material. Deep-level learning strategies are aimed at understanding, distilling meaning, and applying the learning material. After the task has been performed, ideally learners evaluate their performance and attribute the results to causal factors. For students who successfully regulate their learning, this results in conclusions on how to adjust their self-regulatory approach during subsequent learning activities. Therefore, self-regulation is considered a cyclical process that works as a feedback loop; information from previous self-reflection phases is used to adjust current efforts (Zimmerman, 2000).

Mounting evidence shows that students who are able to self-regulate their learning achieve higher learning outcomes (e.g., Dignath, Buettner, & Langfeldt, 2008; Pintrich & De Groot, 1990; Zimmerman & Martinez-Pons, 1986). For long, research on self-regulation has prevalently focused on older students as scholars questioned whether younger students possessed the knowledge and strategies necessary for self-regulated learning (Paris & Newman, 1990; Zimmerman, 1990). Increasingly, however, researchers postulate that students in elementary school are able to regulate their learning (e.g., Perry, 1998; Whitebread *et al.*, 2009). In a similar vein, research focuses more and more on identifying classroom conditions that foster the development of cognitive and metacognitive strategies (Ley & Young, 2001; Paris & Paris, 2001; Perry & VandeKamp, 2000). As part of these classroom conditions, considerable research is devoted to instructional strategies and task characteristics that enable students to develop their self-regulatory strategies. The focus has shifted from decontextualized strategy instruction to modelling the use of strategies in context (Boekaerts & Corno, 2005; Paris & Paris, 2001) and from simple, closed tasks to complex and open-ended activities that support students' autonomy (Paris & Paris, 2001; Perry & VandeKamp, 2000). This changed approach towards instructional activities also calls for a change in assessment activities. In this respect, the importance of aligning instructional activities with assessment activities has been emphasized (Biggs & Tang, 2007). Attempts to enhance strategy use are only achieved when assessment invokes the same strategies as those addressed by the instructional activities set out in the classroom. However, knowledge about the role of assessment in enhancing cognitive and metacognitive strategies is limited.

Assessment for Learning to enhance cognitive and metacognitive strategy use

Numerous scholars in the field of assessment have advocated the integration of assessment practices into instruction to enhance learning (e.g., Birenbaum *et al.*, 2006; Black & Wiliam, 1998b; Klenowski, 2009). The concept of AfL has been introduced as a counterbalance for Assessment *of* Learning (AoL). While the term AfL is popular, there is no unambiguous definition of the term or a set of guidelines of how it should be translated to practice. AfL and AoL are often conceptualized as different concepts; however, they are closely related to each other (Bennett, 2011). AoL is conceptualized as assessment activities that are separated from the curriculum and summative in nature, aimed to record achievement (Birenbaum *et al.*, 2006; Black & Wiliam, 1998b; Harlen & James, 1997). AfL, on the other hand, is characterized as the ongoing process of collecting and interpreting assessment information that takes place in the interaction between teacher, student, and peers (Black, Harrison, Lee, Marshall, & Wiliam, 2004; Klenowski, 2009). Assessment information can be derived from a wide spectrum of activities, such as student work, classroom discussions, or scores on standardized tests. Assessment activities support student learning when the derived information is used to adjust teaching and learning to better fit student needs (Black *et al.*, 2004). This entails that instructional activities are designed to engage students in actions to enhance their learning and increasingly take responsibility for their learning. In the classroom, assessment practices represent a similar cycle as the phases of self-regulated learning at an individual level (Birenbaum *et al.*, 2009). Key processes in enhancing strategy use through AfL are carefully monitoring student progress and providing scaffolds that promote learning (Pat-El, Tillema, Segers, & Vedder, 2013; Wiliam, 2011).

Monitoring activities provide students with information that helps them to understand where they are in their learning (Wiliam, 2011). Feedback is pivotal in this respect (Hattie & Timperley, 2007; Nicol & Macfarlane-Dick, 2006; Sadler, 1998). Teacher feedback provides students an understanding of the gap between their current performance and the learning goals they are aiming for. Eliciting students' reflection on teacher feedback in student–teacher dialogue helps students to acquire metacognitive knowledge on the effectiveness of their learning strategies (Clark, 2012). In addition, facilitating self-assessment enhances students' ability to employ metacognitive strategies. The process of comparing current performance to desired performance generates feedback that helps students in optimizing their learning (Butler & Winne, 1995). After monitoring their learning, students should be able to select and execute strategies to move closer to the goals they are aiming for (Sadler, 1989). Therefore, in order for assessment to drive learning, monitoring student learning is necessary, but not sufficient.

In addition to monitoring progress, assessment activities should provide scaffolds that shed light on what the next steps in learning are (Black & Wiliam, 2009). Students need to have a clear understanding of the learning goals they will be pursuing and the criteria that define good work in order to improve their learning. (Nicol & Macfarlane-Dick, 2006). Therefore, feedback should not only describe where students are in their learning, but also increase their understanding of the steps they need to take to close the gap between their current performance and their goal (Sadler, 1989). In this respect, feedback should provide scaffolds for students that stimulate them to reflect on their learning and should not be prescribing students what they should do. According to Hattie and Timperley (2007), especially information directed towards the level of self-regulation is powerful in eliciting deep processing. Descriptive feedback information on learning, in turn, guides students towards a better use of strategies, such as modifying extant goals, planning learning activities, choosing learning strategies, and adjusting learning strategies when necessary (Butler & Winne, 1995; Clark, 2012; Shute, 2008).

Claims about the impact of AfL on cognitive and metacognitive strategies are supported by studies that show the effectiveness of various elements of AfL. First, in terms of monitoring, research evidenced the impact of formative feedback on metacognitive strategy use. Students who received formative feedback from a tutor and were thereby stimulated to reflect on their learning showed an increase in self-regulation (Van den Boom, Paas, Van Merriënboer, & Van Gog, 2004). Second, regarding scaffolding, studies on the formative use of rubrics have shown that describing both assessment criteria and various quality levels of work supports students in planning their learning and monitoring their work while performing learning activities (Andrade & Du, 2005; Panadero & Jonsson, 2013). In line with this, self-assessment and transparency of assessment criteria affects students' metacognitive skills. Kostons, Van Gog, and Paas (2012) showed that enhancing students' self-assessment skills increased their understanding of the next step in their learning, as was shown in their task-selection accuracy. Third, research on authentic assessment practices in a classroom setting by Perry (1998) compared classrooms where assessment was integrated in ongoing activities and a responsibility of both teacher and student with classrooms where assessment was a separate activity under the responsibility of teachers. Results showed that students showed qualitatively different use of self-regulatory strategies in classrooms where assessment was an ongoing activity with an emphasis on individual progress in both product and process compared with classrooms where assessment was a separate activity that emphasized student differences and the number of correct answers.

The current study

The literature outlined above shows that AfL and the development of cognitive and metacognitive strategies have been considered to be related on theoretical grounds. Remarkably, very few studies have contributed to our understanding of this relation by investigating this relation empirically. This study attempts to fill this gap and provide empirical evidence for this assumed relation. The research question addressed in this study is: 'What is the relation between AfL and students' self-reported cognitive and metacognitive strategy use?' It builds on the hypothesis that the integration of monitoring and scaffolding in classroom practice is positively related to students' use of cognitive and metacognitive strategies.

Method

Participants

The sample consisted of 528 grade four to six students (9- to 12-year-olds) from seven Dutch elementary schools from small towns and rural areas in the south-eastern part of the country. Of the total sample, 271 (51.3%) were girls and 257 (48.7%) were boys. Schools were selected at an introductory meeting about the purpose of a research and development track that aimed to develop AfL by means of portfolio assessment. After this meeting, seven schools were willing to participate in this research. These schools were informed about the nature of the research and the investment it would require from students and teachers.

The participating schools aimed to enhance students' motivation and self-regulated learning by means of AfL. Portfolio assessment was used by the schools to support the integration of AfL in classrooms. Students monitored their development by selecting work for their portfolio. Student work that was added to the portfolio was accompanied by reflection tags that supported students in reflecting on their learning process and the strengths and weaknesses off their work. Furthermore, student and teacher engaged in reflective dialogue about portfolio work and the learning process related to portfolio work. Based on student reflections and teacher feedback, new learning goals were formulated.

Measures

The variables in our model were measured with two questionnaires. The extent to which AfL was embedded in classrooms was measured with the Student Assessment for Learning Questionnaire (SAFL) (Pat-El *et al.*, 2013). Students completed the two scales that measure students' perceptions of monitoring (16 items, $\alpha = .89$) and scaffolding (12 items, $\alpha = .83$). The monitoring scale gauges students' perceptions of the feedback they receive and the extent to which they evaluate their own work. The scaffolding scale measures students' perceptions of processes related to instruction, namely classroom questioning and the clarity of learning goals and criteria. Students' self-regulatory skills were assessed using six scales of the Children's Perceived Use of Self-Regulated Learning Inventory (CP-SRLI) (Vandevelde, Van Keer, & Rosseel, 2013). The scale 'task orientation' (six items, $\rho = .73$) measures activities that students undertake to analyse a learning task prior to executing it, such as activating prior content and metacognitive knowledge and analysing their feelings about the task. The 'planning' scale (five items, $\rho = .54$) gauges the extent to which students plan time and activities and select strategies in order to achieve their learning goals. The scale 'surface learning strategies' (four items, $\rho = .77$) measures strategies to rehearse and memorize information, such as reciting and copying material and repeating the material to be learned aloud. Strategy use to come to a deep and meaningful understanding of the learning material is measured with the scale 'deep-level learning strategies' (10 items, $\rho = .84$). This includes making learning material meaningful by relating it to prior knowledge, looking for examples, separating main issues from side issues, and looking for analogies. Students' strategies in evaluating their learning outcome in terms of completeness and correctness are measured with the scale 'product evaluation' (three items, $\rho = .80$). Lastly, the scale 'process evaluation' (four items, $\rho = .77$) gauges the extent to which students evaluate the process that led to learning outcome, for example by evaluating the effectiveness of their strategy use, their emotions while performing the task, and how they would approach a similar task in the future. Both the SAFL-Q and the CP-SRLI items used a 5-point Likert scale, ranging from 'never' to 'always'. Typical items for each scale are given in Table 1.

Table 1. Overview of the scales used number of items per scale, Cronbach's alphas, and example items

Scale	*N*	α	Example item
SAFL-Q			
Monitoring	14	.87	My teacher inquires which of my assignments went well and which did not
Scaffolding	8	.69	My teacher provides me with hints to help understand the subject matter
CP-SRLI			
Task orientation	4	.74	Before I start my schoolwork, I ask myself: 'Do I know what kind of a task this is?'
Planning	5	.65	Before I start my schoolwork, I decide what to do first and what later
Surface learning strategies	2	.70	When studying, I read or recall everything again and again until I know it by heart
Deep-level learning strategies	8	.78	When studying something new I relate it to things that I already know
Product evaluation	3	.77	After finishing my schoolwork, I go over my answers again
Process evaluation	3	.78	After finishing my schoolwork, I ask myself: 'Will I use a similar approach next time, or should I choose a different approach?'

Note. SAFL-Q; Student Assessment for Learning Questionnaire; CP-SRLI, Children's Perceived Use of Self-Regulated Learning Inventory.

Procedure

The instruments were administered a couple of weeks after the start of the school year during regular class time. Teachers handled the administration of the questionnaires in their classrooms. A period of 6 weeks was determined for this to make sure that the administration of the questionnaires did not interfere too much with regular classroom practice. Both questionnaires provided an example item that illustrated how to work with the Likert scale. All schools and students were assured that responses were treated confidentially.

Analysis

The conceptual model was analysed through structural equation modelling, using EQS 6.1 (Bentler, 1989). First, the measurement model was established. To assess construct validity of the instruments used, confirmatory factor analysis using maximum likelihood estimation was performed on the separate questionnaires. Model modifications were based on the modification indexes, and the number of cross-loadings of items and low factor loadings of items. The Satorra–Bentler scaled statistic (ΔS-Bχ^2) and robust standard errors were evaluated as the data were multivariate non-normally distributed. A combination of goodness-of-fit indices was used to evaluate the fit of the model. We emphasized the comparative fit index (CFI) and the root mean square error of approximation (RMSEA). Bentler (1992) proposes that CFI values greater than .90 indicate an acceptable fit of the model to the data. For the RMSEA, a cut-off score of $<.06$ was employed to indicate a good fit (Hu & Bentler, 1999). Subsequently, the full measurement model that included all the scales used in the current study was evaluated to explore the relations between the indicators and the latent variables. Second, after confirmation of the measurement model, the relations between the latent variables were specified in the structural model. In addition, gender and grade were included as control variables to the structural model to avoid confounding with the variables under study. Previous research has demonstrated gender differences in self-regulated learning strategies (e.g., Vandevelde *et al.*, 2013; Zimmerman & Martinez-Pons, 1990). Furthermore, grade was added as a proxy measure for age as the use of cognitive and metacognitive strategies is strongly related to age (Hübner, Nückles, & Renkl, 2010; Zimmerman & Martinez-Pons, 1990).

Results

Measurement model development

This study builds on the construct validity of questionnaires by Pat-El *et al.* (2013) and Vandevelde *et al.* (2013). As these measures have not been used in the context of Dutch elementary education before, the construct validity of the questionnaires was assessed in the present sample with separate confirmatory factor analyses (CFA) for each questionnaire. The CFA of the original SAFL-Q fitted poorly to the observed data: SB-χ^2 = 830.926; df = 349; p < .001; CFI = .833; RMSEA = .052 (.047, .057). After analysing the parameter estimates, six poorly fitting items were identified because of high error variance and cross-loadings. These items were eliminated from the model after careful analysis. Two items from the monitoring scale were deleted, and four items from the scaffolding scale were deleted. The eliminated items were excluded from subsequent analyses as the topics measured by the deleted items were captured by remaining items and therefore did not cause a substantial loss of information. In addition, correlation was allowed between the residuals of five item pairs loading on the same factors and depicting similar content. The model respecifications resulted in small adjustments to the original questionnaire and a good fit to the data: SB-χ^2 = 289.697; df = 203; p < .001; CFI = .963; RMSEA = .029 (.021, .036).

The CFA of the CP-SRLI scales yielded a moderate fit to the data: SB-χ^2 = 1068.175; df = 449; p < .001; CFI = .863; RMSEA = .053 (.049, .057). After analyses of the items that showed high error variance and cross-loadings, seven items were removed from the model: Two from the scale task orientation, two from the scale surface learning strategies, two from the scale deep-level learning strategies, and one from the scale process evaluation. The elimination of the items was not associated with a loss of concepts measured by the questionnaire. Therefore, the eliminated items were excluded from further analyses. The modifications resulted in a good model fit: SB-χ^2 = 347.166; df = 237; p < .001; CFI = .963; RMSEA = .031 (.024, .038).

Full measurement model

Next, the full measurement model was evaluated that incorporated all the latent constructs under study. The fit statistics of this model indicated an adequate fit to the data: SB-χ^2 = 1182.483; df = 911; p < .001; CFI = .947; RMSEA = .025 (.021, .029). The explained variance of the individual items varied from .160 to .674. An overview of the final scales used, along with the Cronbach's alphas and example items, is presented in Table 1. Table 2 gives an overview of the descriptive statistics and the correlations between the factor scores. In sum, based on the fit indices, the explained variance, and the Cronbach's alphas of the scales, the full measurement model was considered adequate.

Structural model

In the next step, the hypothesized structural relations among the variables were tested in the structural model. This model proved to fit the data well: SB-χ^2 = 1504.011; df = 1031; p < .001; CFI = .913; RMSEA = .032 (.028, .035). The standardized parameter estimates are shown in Figure 1. Results of the structural model show that monitoring positively predicts task orientation (β = .25) and planning (β = .26). Scaffolding positively predicts surface learning strategies (β = .25), deep-level learning strategies (β = .32), and process

Table 2. Means, standard deviations, and intercorrelations

	M	*SD*	1	2	3	4	5	6	7	8
1. Monitoring	3.26	0.67	1.00							
2. Scaffolding	3.84	0.52	.57	1.00						
3. Task orientation	3.17	0.82	.32	.27	1.00					
4. Planning	3.01	0.85	.32	.28	.45	1.00				
5. Surface learning strategies	3.37	1.04	.22	.26	.37	.41	1.00			
6. Deep-level learning strategies	2.88	0.69	.36	.33	.47	.58	.47	1.00		
7. Product evaluation	3.90	0.85	.16	.17	.34	.34	.32	.40	1.00	
8. Process evaluation	3.06	1.00	.23	.24	.50	.46	.21	.51	.47	1.00

Note. All correlations are significant at the .01 level. The range for all scales was 1–5.

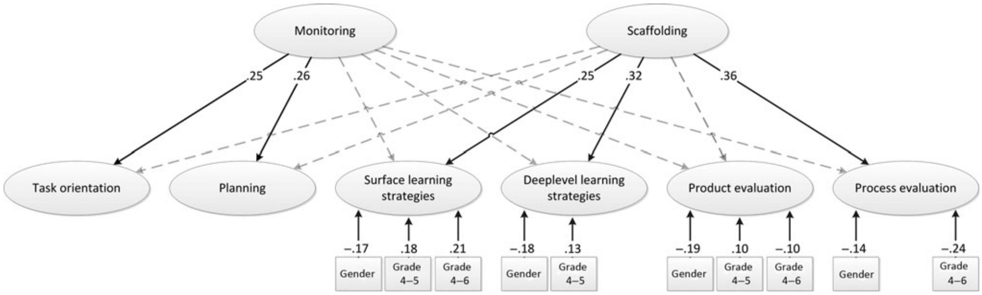

Figure 1. Standardized solutions of the structural model. *Note.* For clarity reasons, correlations among the latent constructs and the error terms are omitted. Path coefficients are significant at the .05 level.

evaluation (β = .36). Product evaluation was neither related to monitoring nor scaffolding.

The control variables, gender and grade, were found to have significant relations with a number of scales. Girls reported higher levels of surface learning strategies (β = −.17), deep-level learning strategies (β = −.18), product evaluation (β = −.19), and process evaluation (β = −.14) than boys did. Our results also reveal significant differences for grade. Grade 6 students surpassed fifth grade students (β = .21) in their use of surface learning strategies, and fifth grade students, in turn, reported more surface learning strategies than fourth grade students (β = .18). Compared with grade 4 students, a greater use of deep-level learning was reported by grade 5 students (β = .13). A complex pattern was found for product evaluation. Grade 5 students displayed more product evaluation than grade 4 students (β = .10). However, sixth grade students reported less product evaluation than fourth grade students (β = −.10). With respect to process evaluation, sixth grade students reported significantly less process evaluation than fourth grade students (β = −.24).

Conclusion and discussion

The main purpose of this article was to unravel the relation between AfL practices and students self-regulated learning (Black & Wiliam, 2009; Clark, 2012; Nicol & Macfarlane-Dick, 2006). AfL has been advocated to provide students with a learning environment in which they can develop their self-regulatory skills. Yet, investigating this relation quantitatively has not received substantial attention among scholars. The current study aimed to fill this gap and increase our understanding of the relation between AfL and elementary school students' cognitive and metacognitive strategies. It was expected that the core AfL principles, monitoring student growth and providing scaffolds to optimize learning, would positively predict students' strategy use. The results reveal that providing students a clear understanding of where they are in their learning (monitoring) predicts students' task orientation and planning activities. Supporting student learning by discussing with students what the next step in their learning is (scaffolding) is positively related to students' use of surface learning strategies, deep-level learning strategies, and process evaluation.

Monitoring activities inform students on their progress and their strengths and weaknesses. Our results show that this information does predict student behaviour in the forethought phase, but not during task execution and self-reflection. Evidently, when feedback gives students insight in their strengths and weaknesses, this raises students' awareness of the learning tasks they will be performing. Whereas feedback information on where students are in their learning can be seen as conditional knowledge in task execution, it is remarkable that this information does not influence students' strategy use and reflection on learning. An explanation for this could be that the feedback received by students on their learning progress may have been of poor quality and therefore not helpful for students during task execution. Therefore, students, especially at this age, may not automatically draw inferences from information on their progress that they can use in steering subsequent learning. They need more support from their teacher to regulate learning.

Scaffolding activities predict students' strategy use during task execution and stimulate them to reflect on their learning process. While empirical evidence on the use of learning strategies is inconsistent (e.g., Choy, O'Grady, & Rotgans, 2012; Dinsmore & Alexander, 2012), it is generally assumed that the use of deep-level learning strategies results in higher quality learning than the use of surface learning strategies (Gijbels, Van de Watering, Dochy, & Van den Bossche, 2005). However, dichotomizing surface learning strategies and deep-level learning strategies entails the danger of giving a too simplified description of strategy use as it suggests that learners are stable in their orientation (Dinsmore & Alexander, 2012). Yet, research suggests that students are not stable in their orientation, but adjust their strategy use to the requirements of the situation and the task (Gijbels, Segers, & Struyf, 2008; Scouller, 1998). This indicates that in certain situations, surface learning strategies are perceived as more effective than deep-level strategies. Students in this study are at an age that they are still developing their understanding of different learning strategies and their insight in when to employ them. The enhanced use of both surface learning strategies and deep learning strategies is a first step for students towards the use of a larger variety of learning strategies and coming to understand when to apply a certain learning strategy.

With respect to gender differences in self-regulated learning strategies, girls reported more frequent use of both surface and deep-level learning strategies. Additionally, higher scores were reported by girls on the scales product evaluation and process evaluation. These results generally confirm findings from previous research (e.g., Ablard & Lipschultz, 1998; Vandevelde *et al.*, 2013; Wall, Higgins, Remedios, Rafferty, & Tiplady, 2013; Zimmerman & Martinez-Pons, 1990). However, Vandevelde *et al.* (2013) did not find any significant differences between boys and girls for process evaluation. Regarding grade level, the relation with the use of metacognitive strategies was not straightforward. Zimmerman and Martinez-Pons (1990) also found the relation between grade and the use of metacognitive strategies to be a complex one. As expected from students' increased experience with the execution of learning strategies, they showed an increase in the use of surface learning strategies from grades 4 to 6. An increase was also found in the use of deep-level learning strategies from grades 4 to 5, but this levelled off. However, the results show a decline in process evaluation from grades 4 to 6. The use of product evaluation increased from grades 4 to 5, but showed a decline after that. A possible explanation for these findings could be that students are not aware any more that they employ these strategies in their learning as they have become automated. Also, this finding could reflect an artefact of the Dutch curriculum. Students in grade 6 reach the end of elementary school, and national large-scale high-stake standardized tests are administered to inform decisions with respect to selection of the level of secondary education. As a result of this, teachers may give less feedback on the use of metacognitive strategies. Likewise, students may be less focused on their learning process, but more on the outcome.

The current study has several limitations that need to be addressed in future research. First, our results are based on cross-sectional data and limited to exploring the relation between AfL and students' use of cognitive and metacognitive strategies. Therefore, the findings do not yield information about the causal relation between the variables. Quasi-experimental research is necessary to confirm the present findings and come to causal inferences about the relation between AfL and students' metacognitive strategies. Furthermore, longitudinal research can increase our understanding of how AfL impacts the development of these strategies over time. Second, the sample of this study was relatively small. Future research should replicate the current findings in a larger sample. A larger sample would also allow to conduct multilevel analyses whereby it would become possible to take into account the classroom and school context in which students are nested.

Third, even though AfL significantly impacts various factors of cognitive and metacognitive strategy use, the explained variance is rather low. Notwithstanding, the findings shed light on one of the factors that affect the complex process of teaching and learning and show that AfL practices play a pivotal role in enhancing learning. We suggest that future research focuses on identifying direct and indirect factors that influence this complex process between AfL and students' cognitive and metacognitive skills.

Fourth, the use of self-reports for measuring cognitive and metacognitive strategies entails limitations. In this respect, Veenman (2011) describes the drawbacks of the use of self-reports in measuring strategy use. Answering questionnaire items requires students to retrieve the strategies they used during task execution from their memory, and this process can suffer from memory distortions. Moreover, the items can prompt students to label their strategy behaviour according to the phrasing of the questionnaire items. Also, the conditions in which the questionnaires are administered can prompt students in labelling their strategy use. For example, when tasks prior to the questionnaire require deep-level learning strategies, students may be inclined to give higher scores on questionnaire items gauging this type of strategy use. Future research should employ different methods for data collection to measure cognitive and metacognitive strategy use. Combining a self-report measure with classroom observations and thinking aloud protocols will yield higher validity.

The results of this study underscore the importance of assessment practice in ceding responsibility to students in taking control of their own learning. Assessment can be a powerful tool to enhance students' use of cognitive and metacognitive strategies. However, scholars in the field of AfL emphasize that its potential relies heavily on teachers' professional competence in assessment (Ayala *et al.*, 2008; Smith, 2011). Hence, the potential of AfL may not be optimally exploited in practice. Indeed, up until now, little attention has been paid to the skills teachers need to acquire to effectively implement AfL in their classrooms (Smith, 2011). Increasing our understanding of the knowledge and skills required by teachers to embed a wide variety of AfL strategies in practice can help us in further understanding the relation between AfL activities and optimizing students' use of cognitive and metacognitive strategies.

References

Ablard, K. E., & Lipschultz, R. E. (1998). Self-regulated learning in high-achieving students: Relations to advanced reasoning, achievement goals and gender. *Journal of Educational Psychology*, *1*, 94–101. doi:10.1037/0022-0663.90.1.94

Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation, 10,* 1–11. Retrieved from http://pareonline.net/getvn.asp?v=10&n=3.

Assessment Reform Group. (2002). *Assessment for learning: 10 principles*. Retrieved from http://assessmentreformgroup.files.wordpress.com/2012/01/10principles_english.pdf

Ayala, C. C., Shavelson, R. J., Ruiz-Primo, M. A., Brandon, P. R., Yin, Y., Furtak, E. M., Tomita, M. K. (2008). From formal embedded assessments to reflective lessons: The development of formative assessment suites. *Applied Measurement in Education*, *21*, 315–334. doi:10.1080/08957340802347787

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy and Practice*, *18*, 5–25. doi:10.1080/0969594X.2010.513678

Bentler, P. M. (1989). *EQS structural equations program manual*. Los Angeles, CA: BM Software.

Bentler, P. M. (1992). On the fit of models to covariances and methodology to the Bulletin. *Psychological Bulletin*, *112*, 400–404. doi:10.1037/0033-2909.112.3.400

Biggs, J. B., & Tang, C. (2007). *Teaching for quality at university* (3rd ed.). Buckingham, UK: Society for Research Into Higher Education & Open University Press.

Birenbaum, M., Breuer, K., Cascallar, E., Dochy, F., Dori, Y., & Ridgway, J. (2006). A learning integrated assessment system. *Educational Research Review*, *1*, 61–67. doi:10.1016/j.edurev.2006.01.001

Birenbaum, M., Kimron, H., Shilton, H., & Shahaf-Barzilay, R. (2009). Cycles of inquiry: Formative assessment in service of learning in classrooms and in school-based professional communities. *Studies in Educational Evaluation*, *35*, 130–149. doi:10.1016/j.stueduc.2010.01.001

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). *Working inside the black box: Assessment for Learning in the classroom*. London, UK: GL Assessment.

Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, *5*, 7–73. doi:10.1080/0969595980050102

Black, P., & Wiliam, D. (1998b). *Inside the Black Box: Raising standards through classroom assessment*. London, UK: GL Assessment.

Black, P., & Wiliam, W. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, *21*, 5–31. doi:10.1007/s11092-008-9068-5

Boekaerts, M. (1997). Self-regulated learning: A new concept embraced by researchers, policy makers, educators, teachers and students. *Learning and Instruction*, *7*, 161–186. doi:10.1016/S0959-4752(96)00015-1

Boekaerts, M. (1999). Self-regulated learning: Where are we today. *International Journal of Educational Research*, *31*, 445–457. doi:10.1016/S0883-0355(99)00014-2

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology: An International Review*, *54*, 199–231. doi:10.1111/j.1464-0597.2005.00205.x

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, *65*, 245–281. doi:10.3102/00346543065003245

Choy, J. L. F., O'Grady, G., & Rotgans, J. I. (2012). Is the Study Process Questionnaire (SPQ) a good predictor of academic achievement? Examining the mediating role of achievement-related classroom behaviours. *Instructional Science*, *40*, 159–172. doi:10.1007/s11251-011-9171-8

Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review*, *24*, 205–249. doi:10.1007/s10648-011-9191-6

Dignath, C., Buettner, G., & Langfeldt, H. P. (2008). How can primary school students learn self-regulated learning strategies most effectively? A meta-analysis on self-regulation training programmes. *Educational Research Review*, *3*, 101–129. doi:10.1016/j.edurev.2008.02.003

Dinsmore, D. L., & Alexander, P. A. (2012). A critical discussion of deep and surface processing: What it means, how it is measured, the role of context, and model specification. *Educational Psychology Review*, *24*, 499–567. doi:10.1007/s10648-012-9198-7

Dunn, K.E., & Mulvenon, S.W. (2009). A critical review of research evidence on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research and Evaluation*, 14. Retrieved from http://www.pareonline.net/pdf/v14n7.pdf

Gijbels, D., Segers, M., & Struyf, E. (2008). Constructivist learning environments and the (im) possibility to change students' perceptions of assessment demands and approaches to learning. *Instructional Science*, *36*, 431–443. doi:10.1007/s11251-008-9064-7

Gijbels, D., Van de Watering, G., Dochy, F., & Van den Bossche, P. (2005). The relationship between students' approaches to learning and the assessment of learning outcomes. *European Journal of Psychology of Education*, *20*, 327–341. doi:10.1007/BF03173560

Harlen, W., & James, M. (1997). Assessment and learning: Differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy and Practice*, *4*, 365–379. doi:10.1080/0969594970040304

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*, 81–112. doi:10.3102/003465430298487

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1–55. doi:10.1080/10705519909540118

Hübner, S., Nückles, M., & Renkl, A. (2010). Writing learning journals: Instructional support to overcome learning-strategy deficits. *Learning and Instruction*, *20*, 18–29. doi:10.1016/j.learninstruc.2008.12.001

Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice*, *30*, 28–37. doi:10.1111/j.1745-3992.2011.00220.x

Klenowski, V. (2009). Assessment for Learning revisited: An Asia-Pacific perspective. *Assessment in Education: Principles, Policy and Practice*, *16*, 263–268. doi:10.1080/09695940903319646

Kostons, D., Van Gog, T., & Paas, F. (2012). Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning. *Learning and Instruction*, *22*, 121–133. doi:10.1016/j.learninstruc.2011.08.004

Ley, K., & Young, D. B. (2001). Instructional principles for self-regulation. *Educational Technology Research and Development*, *49*, 93–103. doi:10.1007/BF02504930

Marton, F., & Säljö, R. (1997). Approaches to learning. In F. Marton, D. Hounsell & N. Entwistle (Eds.), *The experience of learning. Implications for teaching and studying in higher education* (2nd ed.). Edinburgh, UK: Scottish Academic Press.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*, 199–218. doi:10.1080/03075070600572090

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, *9*, 129–144. doi:10.1016/j.edurev.2013.01.002

Paris, S. G., & Newman, R. S. (1990). Developmental aspects of self-regulated learning. *Educational Psychologist*, *25*, 87–102. doi:10.1207/s15326985ep2501_7

Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, *36*, 89–101. doi:10.1207/S15326985EP3602_4

Pat-El, R. J., Tillema, H., Segers, M., & Vedder, P. (2013). Validation of Assessment for Learning questionnaires for teachers and students. *British Journal of Educational Psychology*, *83*, 98–113. doi:10.1111/j.2044-8279.2011.02057.x

Perry, N. E. (1998). Young children's self-regulated learning and contexts that support it. *Journal of Educational Psychology*, *4*, 715–729. doi:10.1037/0022-0663.90.4.715

Perry, N. E., & VandeKamp, K. J. O. (2000). Creating classroom contexts that support young children's development of self-regulated learning. *International Journal of Educational Research*, *33*, 821–843. doi:10.1016/S0883-0355(00)00052-5

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, *16*, 385–407. doi:10.1007/s10648-004-0006-x

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*, 33–40. doi:10.1037/0022-0663.82.1.33

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119–144. doi:10.1007/BF00117714

Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education*, *5*, 77–84. doi:10.1080/0969595980050104
Schunk, D. H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist*, *25*, 71–86. doi:10.1207/s15326985ep2501_6
Scouller, K. (1998). The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Education*, *35*, 453–472. doi:10.1023/A:1003196224280
Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, *78*, 153–189. doi:10.3102/0034654307313795
Smith, K. (2011). Professional development of teachers – A prerequisite for AfL to be successfully implemented in the classroom. *Studies in Educational Evaluation*, *37*, 55–61. doi:10.1016/j.stueduc.2011.03.005
Van den Boom, G., Paas, F., Van Merriënboer, J. J. G., & Van Gog, T. (2004). Reflection prompts and tutor feedback in a web-based learning environment. Effects on students' self-regulated learning competence. *Computers in Human Behavior*, *20*, 551–567. doi:10.1016/j.chb.2003.10.001
Vandevelde, S., Van Keer, H., & Rosseel, Y. (2013). Measuring the complexity of upper primary school children's self-regulated learning: A multi-component approach. *Contemporary Educational Psychology*, *38*, 407–425. doi:10.1016/j.cedpsych.2013.09.002
Veenman, M. V. J. (2011). Alternative assessment of strategy use with self-report instruments: A discussion. *Metacognition Learning*, *6*, 205–211. doi:10.1007/s11409-011-9080-x
Wall, K., Higgins, S., Remedios, R., Rafferty, V., & Tiplady, L. (2013). Comparing analysis frames for visual data sets using pupil views templates to explore perspectives of learning. *Journal of Mixed Methods Research*, *7*, 22–42. doi:10.1177/1558689812450211
Whitebread, D., Coltman, P., Pasternak, D. P., Sangster, C., Grau, V., Bingham, S., Demetriou, D. (2009). The development of two observational tools for assessing metacognition and self-regulated learning in young children. *Metacognition and Learning*, *4*, 63–85. doi:10.1007/s11409-008-9033-1
Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, *37*, 3–14. doi:10.1016/j.stueduc.2011.03.001
Wiliam, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles, Policy and Practice*, *11*, 49–65. doi:10.1080/0969594042000208994
Winne, P. H. (1996). A metacognitive view of individual differences in self-regulated learning. *Learning and Individual Differences*, *8*, 327–353. doi:10.1016/S1041-6080(96)90022-9
Yin, Y., Shavelson, R. J., Ayala, C. C., Ruiz-Primo, M. A., Brandon, P. R., Furtak, E. M., Young, D.B. (2008). On the impact of formative assessment on student motivation, achievement, and conceptual change. *Applied Measurement in Education*, *21*, 335–359. doi:10.1080/08957340802347845
Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, *25*, 3–17. doi:10.1207/s15326985ep2501_2
Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts & P. R. Pintrich (Eds.), *Handbook of self-regulation* (pp. 13–39). San Diego, CA: Academic Press.
Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments and future prospects. *American Educational Research Journal*, *45*, 166–183. doi:10.3102/000283120 7312909
Zimmerman, B. J., & Martinez-Pons, M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, *23*, 614–628. doi:10.3102/00028312023004614
Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, *82*, 51–59. doi:10.1037/0022-0663.82.1.51

Received 8 April 2014; revised version received 3 September 2014